Amazon Lookout for Metrics

- Anomaly detection in time series data
- "Monitor and detect"
- Seems like you set it up and then just keep checking on it to see if there are any anomalies detected, but don't change anything with regard to the model itself

Amazon Fraud Detector

- Detects fraudulent activates
- Analyzes events/transactions
- Geared toward time-type events like transactions, but doesn't need to be transactional or time based
- Allows for exploration of feature importance within the underlying model

The "model score" on the left is essentially the threshold for 1/0. You can interact with it and click different thresholds and see how it changes the confusion matrix
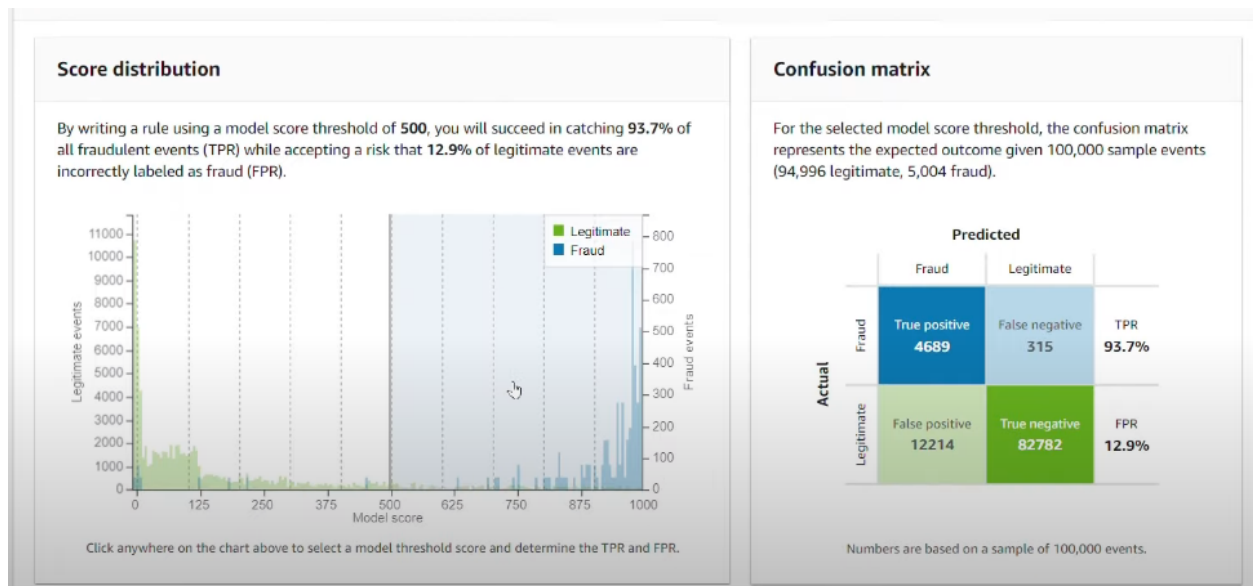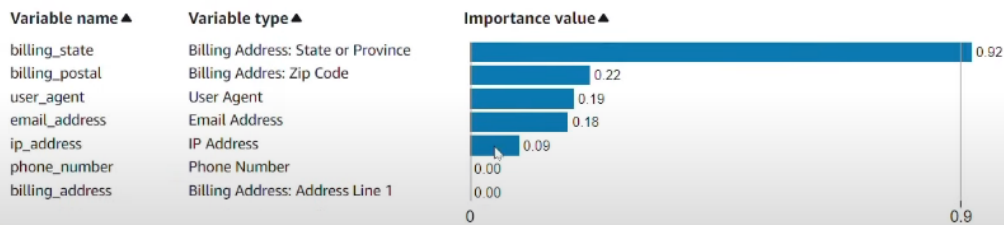
ROC Curve
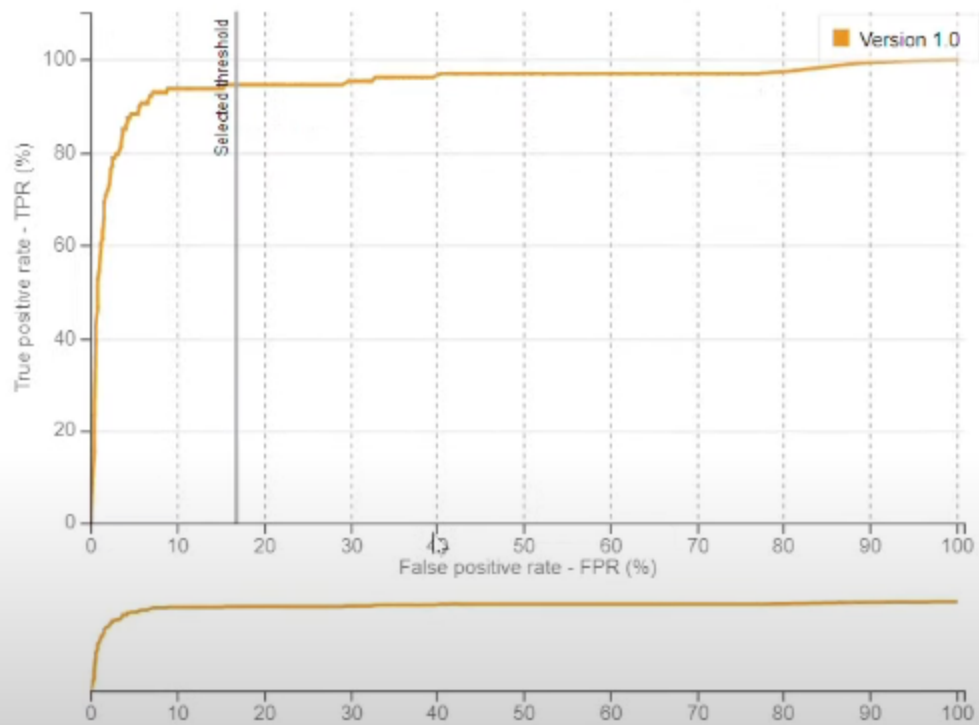
Model variable importance

Variable importance gives you an understanding of how different variables are contributing to your model's performance. The chart below lists variables in the order of its importance to the model, indicated by the number. Variables of AGGREGATE variable type are a combination of multiple variables that are enriched and have an aggregated importance value. A variable (raw or aggregate) with a much higher number relative to the rest could indicate that the model might be overfitting on it, while variables with relatively lowest numbers could just be noise. For more information, see documentation ⧉

Variable name ▲	Variable type ▲	Importance value ▲
billing_state	Billing Address: State or Province	0.92
billing_postal	Billing Addres: Zip Code	0.22
user_agent	User Agent	0.19
email_address	Email Address	0.18
ip_address	IP Address	0.09
phone_number	Phone Number	0.00
billing_address	Billing Address: Address Line 1	0.00

Use the table below to determine which model threshold you should use when writing rules to evaluate events. Choose the model threshold based on the optimal true positive rate (TPR), false positive rate (FPR), and precision scores for your use case.

✓ **How should I interpret this performance data?**   ✕
  - The overall performance of this model is **very high** with an AUC (area under the curve) score of **0.95**. AUC summarizes the true positive rate (TPR) and false positive rate (FPR) across all possible model thresholds. A model with no predictive power will have an AUC of 0.5, whereas a perfect model will have a score of 1.0.
  - Based on the fifth row in the table below, by accepting a risk that **4%** of legitimate events are incorrectly labeled as fraud (FPR), you will succeed in catching **87%** of all fraudulent events (TPR) by writing a rule using a model score threshold of **810**. If you send events with model scores *greater than* the **810** score threshold for manual investigation, **53%** of those events would be fraudulent (precision).
  - Refer to the table to decide which model score threshold is best for your use case.

  If you would like to discuss your model's performance with the Amazon Fraud Detector team, contact us. ↗

False positive rate (FPR) ▲	True positive rate (TPR) ▽	Precision ▽	Model threshold ▽
0%	10%	87%	995
1%	61%	75%	945
2%	73%	65%	900
3%	80%	57%	850
4%	87%	53%	810
5%	88%	48%	765
6%	91%	44%	740
7%	93%	41%	695
8%	93%	38%	655
9%	94%	35%	615
10%	94%	33%	565
20%	94%	20%	345

Fraud Detection with Graph Neural Networks

Graph

User features

GNN Layer

GNN Layer

Node embeddings

Classifier

At a high level, the